

IOWA STATE UNIVERSITY

Digital Repository

Management Publications

Management

6-19-2020

Information flows and topic modeling in corporate governance

Jeffrey D. Kushkowski

Iowa State University, kushkows@iastate.edu

Charles B. Shrader

Iowa State University, cshrader@iastate.edu

Marc H. Anderson

Iowa State University, mha@iastate.edu

Robert E. White

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/management_pubs



Part of the [Business Administration, Management, and Operations Commons](#), [Collection Development and Management Commons](#), [Corporate Finance Commons](#), [Economics Commons](#), and the [Management Sciences and Quantitative Methods Commons](#)

The complete bibliographic information for this item can be found at https://lib.dr.iastate.edu/management_pubs/57. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

This Article is brought to you for free and open access by the Management at Iowa State University Digital Repository. It has been accepted for inclusion in Management Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Information flows and topic modeling in corporate governance

Abstract

Purpose – Multiple disciplines such as finance, management and economics have contributed to governance research over time. However, the full intellectual structure of the governance “field” including the exchange of knowledge across disciplines and the large variety of governance topics remains to be uncovered. To appreciate the breadth of corporate governance research, it is necessary to understand the disciplinary sources from which the research stems. This manuscript focuses on the interdisciplinary underpinnings of corporate governance research.

Design/methodology/approach – This paper employs bibliometric analysis to trace the evolution of corporate governance using articles included in the ISI Web of Science database between 1990 and 2015. Journals included in these categories encompass a full range of business disciplines and provide evidence of the multi-disciplinary nature of corporate governance. It also uncovers the topics treated by disciplines under the governance umbrella using a machine learning method called latent Dirichlet allocation (LDA).

Findings – Corporate governance research deals with a number of strategy-related topics. Unlike strategy topics that reside in a single discipline, corporate governance crosses disciplinary boundaries and includes contributions from accounting, finance, economics, law and management. Our analysis shows that over 80% of corporate governance articles come from outside the field of management. Our LDA solution indicates that the major topics in governance research include corporate governance theory, control of family firms, executive compensation and audit committees.

Originality/value – The results illustrate that corporate governance is far more interdisciplinary than previously thought. This is an important insight for corporate governance academics and may lead to collaborative research. More importantly, this research illustrates the usefulness of LDA for investigating interdisciplinary fields. This method is easily transferable to other interdisciplinary fields and it provides a powerful alternative to existing bibliometric methods. We suggest a number of topic areas within library and information science where this method may be applied, including collection development, support for interdisciplinary faculty and basic research into emerging interdisciplinary areas.

Keywords

Corporate governance, Bibliometrics, Citation analysis, Latent Dirichlet allocation, Agency theory, Director, Interdisciplinary

Disciplines

Business Administration, Management, and Operations | Collection Development and Management | Corporate Finance | Economics | Library and Information Science | Management Sciences and Quantitative Methods

Comments

This article is published as Kushkowsky, J.D., Shrader, C.B., Anderson, M.H., White, R.E., Information flows and topic modeling in corporate governance. *Journal of Documentation*, June 19, 2020; Doi: doi: [10.1108/JD-10-2019-0207](https://doi.org/10.1108/JD-10-2019-0207).

Creative Commons License



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This article is available at Iowa State University Digital Repository: https://lib.dr.iastate.edu/management_pubs/57

Information flows and topic modeling in corporate governance

Information
flows and topic
modeling

Jeffrey D. Kushkowski

University Library, Iowa State University, Ames, Iowa, USA, and

Charles B. Shrader, Marc H. Anderson and Robert E. White

Department of Management, Ivy College of Business, Iowa State University, Ames, Iowa, USA

Received 17 October 2019
Revised 5 February 2020
Accepted 6 February 2020

Abstract

Purpose – Multiple disciplines such as finance, management and economics have contributed to governance research over time. However, the full intellectual structure of the governance “field” including the exchange of knowledge across disciplines and the large variety of governance topics remains to be uncovered. To appreciate the breadth of corporate governance research, it is necessary to understand the disciplinary sources from which the research stems. This manuscript focuses on the interdisciplinary underpinnings of corporate governance research.

Design/methodology/approach – This paper employs bibliometric analysis to trace the evolution of corporate governance using articles included in the ISI Web of Science database between 1990 and 2015. Journals included in these categories encompass a full range of business disciplines and provide evidence of the multi-disciplinary nature of corporate governance. It also uncovers the topics treated by disciplines under the governance umbrella using a machine learning method called latent Dirichlet allocation (LDA).

Findings – Corporate governance research deals with a number of strategy-related topics. Unlike strategy topics that reside in a single discipline, corporate governance crosses disciplinary boundaries and includes contributions from accounting, finance, economics, law and management. Our analysis shows that over 80% of corporate governance articles come from outside the field of management. Our LDA solution indicates that the major topics in governance research include corporate governance theory, control of family firms, executive compensation and audit committees.

Originality/value – The results illustrate that corporate governance is far more interdisciplinary than previously thought. This is an important insight for corporate governance academics and may lead to collaborative research. More importantly, this research illustrates the usefulness of LDA for investigating interdisciplinary fields. This method is easily transferable to other interdisciplinary fields and it provides a powerful alternative to existing bibliometric methods. We suggest a number of topic areas within library and information science where this method may be applied, including collection development, support for interdisciplinary faculty and basic research into emerging interdisciplinary areas.

Keywords Corporate governance, Bibliometrics, Citation analysis, Latent Dirichlet allocation, Agency theory, Director, Interdisciplinary

Paper type Research paper

© Jeffrey D. Kushkowski, Charles B. Shrader, Marc H. Anderson and Robert E. White. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial & non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at: <http://creativecommons.org/licences/by/4.0/legalcode>.

The authors wish to thank Florence Honore of the University of Wisconsin for help with the early stages of this project and feedback on the manuscript; Kyle Hansen, Manaswi Podduturi and Hope Scheffert of Kingland Systems for their help with the latent Dirichlet allocation analysis; and, Kelly Moore for assistance with data collection. Thanks to the Iowa State University Ivy College of Business and the University Library for financial support of the open access fees.



Journal of Documentation
Emerald Publishing Limited
0022-0418
DOI 10.1108/JD-10-2019-0207

1. Introduction

Corporate governance is the area of business that deals with the relationships between boards of directors, company management, shareholders, and other stakeholders. It describes the relationships between these groups that allow corporate entities to thrive and deals with a wide range of economic, financial, and legal topics. Interest in governance topics has grown substantially since 1990 (Turnbull, 1997; Durisin and Puzone, 2009; Kushkowski and Shrader, 2013). From a relatively modest beginning, a great deal of research-based knowledge now emanates from the sub-disciplines of this interdisciplinary field of study (Turnbull, 1997; Yoshikawa and Rasheed, 2009). Research in corporate governance exists in the business fields of accounting (Keasey *et al.*, 2005), finance and law (Shleifer and Vishny, 1997), strategy (Pugliese *et al.*, 2009) and behavioral theory (Van Ees *et al.*, 2009).

Despite the contributions to the literature from disparate business disciplines, management researchers contend that corporate governance has its basis in the management literature (Chen and Chan, 2010). This paper describes the multi-disciplinary underpinnings of corporate governance research and traces the evolution of corporate governance research using articles included in the ISI *Web of Science* database between 1990 and 2015. Journals included in these categories encompass a range of business disciplines. Our results show that corporate governance research has rich, multidisciplinary roots and will hopefully spur greater interdisciplinary collaboration between governance researchers.

More importantly, for LIS researchers, the methods from this paper can be used as a supplement to existing bibliometric methods. There are a number of areas within LIS where the method we use could be applied by researchers and librarians, including collection development and management, support for interdisciplinary research, information seeking behavior, and applied and basic research into emerging and established interdisciplinary areas.

Bibliometric analyses are conducted for several reasons. They define a field and its researchers (Samiee and Chabowski, 2012), they provide a retrospective look at a given journal or set of journals (Koseoglu, 2016), they explain the evolution of new research areas (Acedo and Casillas, 2005), and they identify the intellectual structure of a field (Culnan, 1987).

The citation metrics of a field or discipline are important because they provide context for the development of a field. Disciplines are not static entities; they grow, develop, and change in response to the research done in the discipline. The heterogeneous nature of disciplinary content within corporate governance research makes determining the intellectual structure of the field more complicated.

The discovery and explanation found in intellectual structure articles is analogous to review articles in intent – to describe the present state of a field. The difference is that research on intellectual structure rely on the statistical analysis of bibliometric data rather than a non-empirical reflection on the research literature (Župič and Čater, 2015). To appreciate the scope of corporate governance research, it is necessary to examine and understand the varied disciplinary sources in which corporate governance research is located.

This paper investigates this progression by using the bibliometric methods outlined below. This review expands on previous bibliometric efforts by considering journal articles published in the governance field from 1990 to 2015 and all the journal citations contained in these papers. This research also identifies how the citations are related to the primary sub-disciplines of governance research (accounting, economics, finance, law, management), and present the first comprehensive analysis of the evolution of the governance field.

In addition to traditional bibliometric analysis, this research employs latent Dirichlet allocation (LDA), an objective and powerful form of analysis that allows a computer to “learn” which topics exist in a large body of text. The LDA algorithm allocates words to topics based on their similarity as derived from a set of words called a corpus. LDA is a machine learning algorithm used to explore large amounts of textual data to discern which topics emerge and how those topics are related. The computer learns what words coincide with each other and

uses probabilistic topic modeling to estimate the likelihood that words are grouped in similar topic areas (Župić and Čater, 2015) and how topics are grouped in documents (individual transcripts). Our corpus consists of the set of articles retrieved in a search of topics related to corporate governance.

2. Citation analysis

The application of quantitative analysis to a body of academic citations is often referred to as bibliometric analysis or bibliometrics (Župić and Čater, 2015). In bibliometric analysis, counts are performed over a period of time to establish the total number of citations received by a particular article or author from a set of documents existing in the literature. The underlying assumption is that highly-cited authors and articles have greater impact on the development of a discipline than do those cited less frequently (Culnan, 1986). Citation analysis, therefore, offers a means by which academics interested in corporate governance can measure disciplinary contributions to a field.

Bibliometric analysis can be classified into two basic types: (1) citation analysis, which involves the counting of citations and (2) network analysis, which examines the relationships among citations and which includes co-citation analysis (Georgi *et al.*, 2010) and related methods. Our analysis in this paper uses citation analysis. Citation counts of highly cited papers and authors are often used to indicate scholarly influence of papers and authors and the intellectual structure of academic fields. Citation analysis can also be used to evaluate the performance of individual authors, journals, and institutions (Kushkowsky and Shrader, 2013). Citation and network analysis frequently uses authors, articles, or journals as source data for analyzing relationships between and within disciplines (White and McCain, 1998). For example, Ramos-Rodriguez and Ruiz-Navarro (2004) used bibliometric analysis to track the intellectual structure of strategic management research, including identifying the most influential books, articles, and authors.

Bibliometric papers examining the intellectual structure of business disciplines range in scope from broad overviews of management literature (e.g. Podsakoff *et al.*, 2008) to granular studies of specific topics, such as Shiau and Dwivedi's (2013) study of the e-commerce literature. Regardless of their scope, intellectual structure papers provide a framework for exploring trends in a discipline. Citation or bibliometric analysis has been used to examine the intellectual structure of many business fields including: business ethics (e.g. Calabretta *et al.*, 2011), entrepreneurship (e.g. Kushkowsky, 2012), human resources (e.g. Fernandez-Alles and Ramos-Rodriguez, 2009), information systems (e.g. Culnan, 1987), logistics/supply chain (e.g. Georgi *et al.*, 2013), management and religion (e.g. Gundolf and Filser, 2013), marketing (e.g. Hult, 2015), product innovation (e.g. Durisin *et al.*, 2010; Shafique, 2013), risk management (e.g. Chiang and Yang, 2012) and strategic management research (e.g. Ramos-Rodriguez and Ruiz-Navarro, 2004; Azar, and Brock, 2008; Nerur *et al.*, 2008).

At present, there is no comprehensive citation analysis describing the intellectual structure of governance research in the extant literature. Filatotchev and Boyd (2009) specifically call for interdisciplinary aspects of governance research in their guest editorial for a special review edition of *Corporate Governance: An International Review*. Durisin and Puzone (2009) come closest to a comprehensive analysis in their review and co-citation analysis of one thousand publications. They rank the most highly cited articles and identify seven research themes for articles published in the journal *Corporate Governance: An International Review*: review articles, agency theory, governance structures, board characteristics, stock ownership and performance, codes of practices, and monitoring executive compensation. Their research also looked at papers published in general management and business journals and produced a similar categorization.

Durisin and Puzone (2009) use author co-citation analysis as a method for showing relationships between topics in the field. Our paper takes a different approach by defining the

essence of corporate governance as a set of topics. By searching topically rather than by journal, this research captures a broader collection of journals where corporate governance research is published.

Previous citation analyses of the governance literature point to the need for enhancing both the scope of the methodology used and the theoretical perspectives of the analysis. For example, both [Cheng and Chan \(2010\)](#) and [Saggese *et al.* \(2016\)](#) argue to broaden the perspective of analysis to include aspects beyond the analysis of articles and the topic of financial control of organizations. [Acedo *et al.* \(2006\)](#) and [Shafique \(2013\)](#) mapped the intellectual structure of the internal capabilities of firms while calling for more interdisciplinary analyses. Likewise, [Turnbull \(1997\)](#), [Yoshikawa and Rasheed \(2009\)](#) and [Tihanyi *et al.* \(2014\)](#) argued for analyses to go beyond the traditional agency conflicts between shareholders and managers and include considerations of the international aspects of governance.

In response to the above, our analysis is a more comprehensive and interdisciplinary effort. It includes international and business functional areas as well as legal studies, and uses a much larger and more comprehensive key-word driven set of governance articles and considers a longitudinal time frame. Our analysis is based on the notion that widely cited authors and papers are deemed to have exerted greater influence on a particular field than less-cited papers ([Sharplin and Mabry, 1985](#); [Culnan, 1986](#)). As a set of ideas grow into a discipline, the rigor of related articles and the focus of journals coalesce into clear patterns and networks ([Kushkowski and Shrader, 2013](#)). Along with this, our paper presumes that highly cited articles, authors, and journals define the development of a field by indicating the key intellectual roots of a discipline ([White and McCain, 1998](#)). Our paper clearly identifies the most highly cited journals in the corporate governance field and maps the metrics of the field for future scholarly research.

3. Latent Dirichlet allocation (LDA), topic modeling and words

I like good strong words that mean something. . . -

Louisa May Alcott, *Little Women*

Current advances in machine learning have led to topic modeling methods that can be used as an exploratory technique to uncover hidden or latent relationships present in large data sets ([Blei, 2012](#)). It is a method of discovery that [Moro *et al.* \(2015\)](#) claim is useful for conducting progressive and relevant research in any disciplinary field. Implementations of LDA and its variants are available from data analytics firms, academic software repositories, or from open source repositories such as Github.

The main difference between these methods and traditional statistical methods is the absence of *a priori* assumptions about relationships present in the data. LDA does not require researchers to make assumptions in terms of how the information in the raw data is sorted into silos (i.e. topics). LDA is an unstructured machine-learning algorithm that uses probabilistic topic modeling to estimate the likelihood that words are grouped in similar topic areas ([Župič and Čater, 2015](#)), and the likelihood that topics are grouped in documents (individual interview transcripts). Observations in a large set of data (called a corpus) are explained by similarities in otherwise unclassified groups ([Blei *et al.*, 2003](#)). Because LDA extracts patterns without *a priori* assumptions, it is not designed to answer specific pre-formulated hypotheses ([Schwab and Zhang, 2018](#)). In this way, LDA opens up large data sets to unrestrained discovery.

As an unstructured machine-learning algorithm, LDA requires a researcher to input textual data to the algorithm which uses joint probability distributions to discover the “hidden structure” of topics within documents and words within topics. The richness of LDA is that it recognizes that there can be many probabilistic topics and that words are

independently distributed among topics (Sugimoto *et al.*, 2011). LDA is used in this study to look for the most mentioned words in a large set of articles dealing with the general topic of corporate governance.

In LDA, a “word” is the basic unit of discrete data (Blei *et al.*, 2003). A “document” or article is a sequence of words, and a “corpus” is a collection of articles. The basic unit of modeling is the topic, which is a distribution of words over the entire set of words in a corpus (Sievert and Shirley, 2014). In effect, LDA helps discover underlying themes in a set of data by generating key words. Words are allocated to topics in the analysis.

The authors worked jointly with technical experts from Kingland Systems, a data analytics firm that provides information technology services to help firms manage compliance with regulations and risk and whose clients include some of the world’s largest banks, financial services firms, and insurance companies. After receiving the project data, Kingland Systems did the data analysis, and the authors collaborated with them to interpret the LDA results. Staff from Kingland provided the technical expertise, and the authors served as the subject domain experts. Each topic is a cluster of words specified by machine learning algorithm. Unimportant words (e.g. a, and, are, is, the) are often ignored or reduced in the analysis. Rather, the important key words, words that are idiosyncratic and exclusive to a topic, are identified and emphasized.

LDA does not create new words or concepts as do factor analysis and content analysis; rather, topics in a set of data are represented by existing keywords (Moro *et al.*, 2015). In a well-formed topic model, certain topics will generate words from one conceptual area more than from another. Topics are based on probabilities assigned by the LDA software. Blei (2012, p. 78) summarizes it as “reversing the generative process – what is the hidden structure that likely generated the observed collection?” where the observed collection is the words in the transcripts.

In LDA, the research administrator determines the number of topics, which can be decided in several ways (Chen and Wang, 2018). Too many topics dilute the meaning of each topic and too few fail to adequately separate ideas and words from each other. The appropriate number of topics for our dataset was determined by comparing the intra-topic similarity with inter-topic dissimilarity. Analyses specifying five, ten, fifteen, and twenty topic solutions were run, and the optimum number of topics was determined to be twenty as this maximized the difference between topics (Chen and Wang, 2018). Words in topics were generated by determining the mix between the probability and relevance of words belonging to topics (Sievert and Shirley, 2014).

The LDA relevance metric sets the weight given to the probability of a word belonging to a topic (Sievert and Shirley, 2014). Our LDA model produced a list of the most relevant words for the twenty topics, where relevance of word w to topic k given λ is defined as:

$$r(w, k/\lambda) = \lambda \log(\varphi_{KW}) + 1 - \lambda \log(\varphi_{KW}/p_W)$$

where φ_{kw} is the probability that word w belongs to topic k and p_w is the probability of word w being in the corpus. Lambda, which serves to define relevance, can be set to any value between and including 0 to 1. A lambda of 0 would equate relevance to exclusivity; the most frequent words have the same probability of being within a particular topic and appearing in the corpus. By contrast, setting lambda to 1 means a word is relevant if it appears in the corpus regardless of which or how many topics it appears in. As recommended by Sievert and Shirley (2014), the developers of the relevance measure, a lambda of 0.6 is used for these results.

Dyer *et al.* (2017) use LDA to ascertain exactly “what is being said” in corporate reports. Župič and Čater (2015, p. 457) state that LDA holds tremendous potential for “expanding the scope of mapping the management and organization domain.” In order to make wide application of LDA available in organization research, Župič and Čater indicate that

management scholars can use available software themselves or work with information scientists on research projects using LDA. The analyses for this research was completed with technical experts from Kingland Systems, a data analytics firm headquartered in Clear Lake, Iowa, which provides information technology and financial services to help firms manage compliance with regulations and risk. For example, Kingland provides their clients full-text analysis of formal corporate reports and compliance documents. Their clients include some of the world's largest banks, financial services firms, and insurance companies. The authors worked directly with Kingland analysts, providing them the raw data, or corpus. Kingland produced the LDA results and collaborated with the authors to interpret the results. The Kingland experts directly helped the authors ascertain the number of LDA topics to best analyze the corpus and they offered advice on the relevance metric to best identify topic words.

The LDA analysis was done using LDAvis software to analyze our data (Sievert and Shirley, 2014), an extension of LDA which provides LDA-based data visualization. LDAvis allows us to examine how different words contribute to the meaning of each topic, how different topics relate to each other, and the prevalence of each topic. The accompanying tables were created from the LDAvis results.

4. Data collection methods

4.1 Common data for citation analysis and LDA

Initial data for this study was harvested from ISI's *Web of Science* database. Journal articles containing the keywords "corporate governance," "agency theory," "director," "market for corporate control," "ownership structure" or "executive compensation" were selected to capture essential elements of corporate governance. A decision was made early in the project to be economical in the choice of search terms and focus on the essence of corporate governance research. A broader keyword search would expand the results set but increase the likelihood of extraneous results.

The search was performed in the Web of Science Social Science Citation Index categories of Business Finance, Business, Management, Economics, and Law to capture multidisciplinary articles on the search topics. The base bibliographic data used for the citation analysis and LDA for this study includes bibliographic information, abstracts and cited references for 10,532 articles from ISI's *Web of Science* database published between 1990 and 2015. The longitudinal data allows for an examination of trends in corporate governance scholarship over time. The twenty-five-year timespan for the data coincides with the growth of corporate governance as a research field and provides convenient data cohorts. The number of articles retrieved by our search for the years 2016 to 2018 were 1,044, 1,162, and 1,243, respectively, demonstrating that corporate governance continues to be a vigorous research area.

In addition, the dataset for LDA included 1,053 full-text articles chosen randomly from the *Web of Science* results. The full-text articles provided a corpus of text in addition to the bibliographic article data from *Web of Science*. A list of the full text sources used in this project is included as part of the supplemental data available for this research.

4.2 Citation analysis dataset

The citation analysis dataset includes the title, journal title, the year of publication, and subject classifications of the journal. These classifications are the categories assigned in Anne-Wil Harzing's *Journal Quality List* (Harzing, 2015) or subject classifications assigned to journals by the Library of Congress and then mapped to a Harzing discipline (for journals not included in Harzing).

Classifying articles by discipline provides insights into disciplinary contributions to the literature of corporate governance (Kushkowski and Shrader, 2013). The article subject code is used as a proxy for the discipline of the journal. For example, the *Strategic Management Journal* subject code is “management.” Subject analysis leads to a more nuanced understanding of the development of corporate governance as a field.

The articles downloaded from *Web of Science* included the full bibliographic data and cited references. Over 656,000 cited references were included in the set of articles analyzed to explain trends in authorship. To focus our analysis on academic literature, the following categories of materials were excluded from the analysis: books, conference proceedings, foreign language materials, newspapers, unpublished articles and unidentifiable citations resulting in a set of 415,304 cited references. In the analysis of journals that are most highly cited, publications with fewer than 100 cited references were excluded.

5. Results

Analysis of the database of articles traces disciplinary changes in corporate governance articles by looking at which journals publish the research. Analysis of the citation data provides insight into which journals have been most frequently cited over time. Analysis of the citation data focused on three areas. First, the subject distribution of articles in our dataset and how the subject distribution has changed over time. Second, the subject distribution of the cited references which allows us to make observations about how corporate governance literature is distributed in the academy over time. Third, the journal distribution and which journals contribute the most to corporate governance research.

5.1 Subject analysis of articles

One of the ways to describe the changes in corporate governance research is to look at the subject distribution of articles downloaded from *Web of Science*. The subject categories assigned to the articles provide a straightforward way of describing changes in the disciplinary makeup of corporate governance. The results report totals for the Harzing subject areas of accounting, corporate governance, economics, finance, law, management, and organizational behavior—categories which include 85% of the articles downloaded in our sample. The category of “Other” includes the remaining 15% or articles that include the Harzing categories of business history, entrepreneurship, international business, innovation, marketing, management information systems, multidisciplinary, operations research, political science, public service management, psychology, sociology and tourism.

An easy way to gauge the growth of corporate governance scholarship is to look at yearly publishing output. Figure 1 shows the number of articles produced each year by subject area.

One trend to notice is that the growth of corporate governance scholarship was relatively stable between 1990 and 2004. Beginning with 2004, the total number of articles on corporate governance topics increased steadily with finance journals publishing more than other areas. This may be the result of the academic research in the impact of Sarbanes–Oxley Act which was passed in 2002, though the Act was not included in the search terms. Articles in journals dedicated to corporate governance, however, only began to appear in 2000, even though the main academic title, *Corporate Governance: An International Review*, started publication in 1993. The lack of explicit corporate governance articles prior to 2000 may be due to articles in the journal and the search terms not matching up. The total number of articles has increased each year and there is a broad distribution of articles across subject areas.

Another way to view disciplinary output is to look at the percentage of articles in each subject area over time. Figure 2 shows the percentage of the total output of governance articles that come from each field.

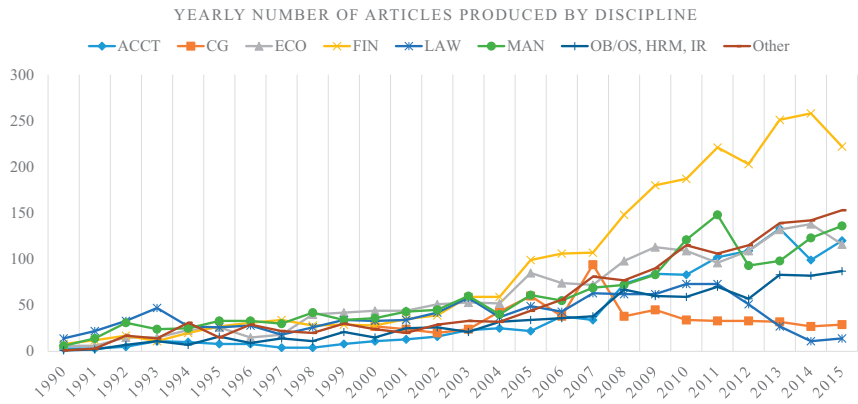


Figure 1.
Yearly number of
articles produced by
discipline (1990–2015)

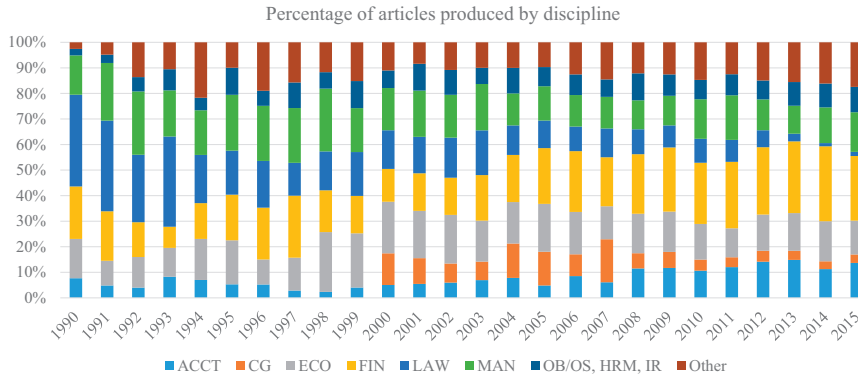
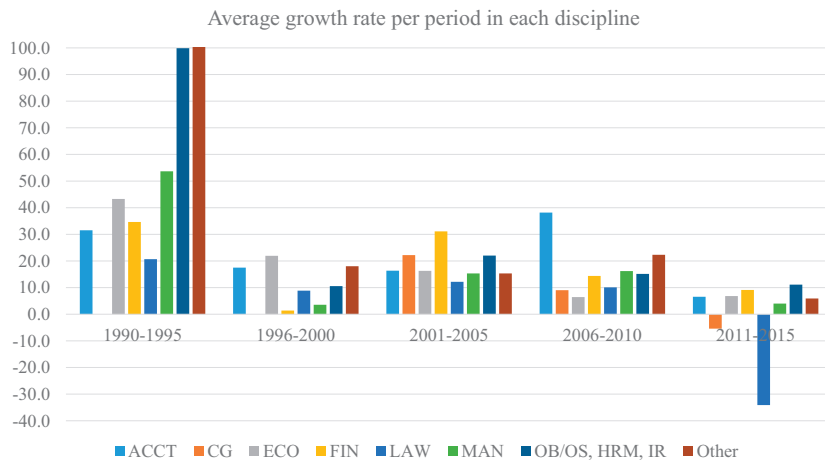


Figure 2.
Percentage of articles
produced by discipline
(1990–2015)

As expected, finance has contributed a large share since the beginning and has been able to maintain it over time. Management has had a constant but smaller share of articles, an indication that corporate governance research is appearing in other disciplines. The lack of corporate governance articles in the results is explained by the fact that there are few pure corporate governance journals.

The number of articles published and the share of articles published by subject area tells part of the story behind the growth in corporate governance research. Figure 3 illustrates the growth rates of disciplines for subject areas in five year cohorts.

We divided our data into five periods, the first one of six years (1990–1995) and the following of five years each. The yearly growth is not as meaningful as the variation might be due to peculiar changes in journals and to the delay in the review process. Of the single disciplines, organizational behavior grew the most during the first period with 100%. The “other” category grew of 143% (capped at 100% in the figure) mainly due to growth in corporate governance research in the myriad “other” categories. During the next period from 1996 to 2000, the overall growth was modest with the highest rate being reached by economics with around 21.9%. The third period from 2001 to 2005 saw a high average growth across all disciplines with a high rate of corporate governance of 22.2%. In the fourth period from 2006 to 2010, accounting had the highest rate with 38.6%. This might be due to the



Note(s): *The Other category grew of 143% (capped at 100% in the figure) during the 1990-1995 period

Figure 3.
Average growth rate of
articles by discipline by
quintile

research on the implementation and consequences of Sarbanes–Oxley, which was enacted in 2002. In the last period, law experienced a dramatic decline of 34.1% while all the other disciplines experienced a small growth rate.

5.2 Subject analysis of cited references

Governance researchers, mostly management academics, assume that a bulk of governance researcher takes place solely in management (Chen and Chang, 2010). Our results illustrate that this is not the case. The subject breakdown shown in Table 1 for cited references in articles downloaded from the *ISI Web of Science* database is as follows: 25% finance, 18% management, 18% economics, 14% for law, 11% for accounting and 13% for a collection of 12 other disciplines, and less than 2% corporate governance. The low number of cited references for corporate governance is a result of only two journals categorized in this subject area by Harzing.

Our analysis suggests that corporate governance is far more interdisciplinary than previously thought. Management literature, in fact, accounts for less than 20% of the cited references for the articles in our sample.

Our research also included an analysis of the journals appearing in the cited references in our journal sample. Journals are included if they include 100 or more cited references, which yields a total of 289 journals and 273,893 cited references. Table 2 lists the top 50 journals ranked by number of cited references. The titles in this list account for 62% (169,238) of the cited references.

Information about journals that contribute cited references to corporate governance research provides clues about the disciplinary underpinnings of the field. Table 2 shows that the top two journals – *Journal of Finance* and *Journal of Financial Economics* – account for more than twice as many cited references as the next two titles which are in management – the *Academy of Management Journal* and *Strategic Management Journal*. In the top 50 titles, 58% of the cited references come from the fields of accounting, economics, or finance, while only 23% come from the management literature. The conclusion from this analysis is that cited references in corporate governance research are predominantly from fields other than

Field/discipline	Titles	Cited ref.	% Cited ref.
Accounting	27	29,098	10.62%
Business history	2	445	0.16%
Corp governance	2	4,071	1.49%
Economics	61	49,608	18.11%
Entrepreneurship	6	4,260	1.56%
Finance	26	69,020	25.20%
Innovation	2	360	0.13%
International business	6	3,793	1.38%
Law	71	37,976	13.87%
Management	30	49,914	18.22%
Management information systems	3	482	0.18%
Marketing	12	4,575	1.67%
Operations research	7	2,915	1.06%
Organizational behavior	16	10,838	3.96%
Political science	1	165	0.06%
Psychology	8	2,720	0.99%
Public sector management	3	405	0.15%
Sociology	6	3,248	1.19%
Grand Total	289	273,893	100.00%

Table 1.
Cited references by discipline

Note(s):
(1) Used Harzing’s *Journal Quality List* to assign subjects to journals
(2) This table shows totals for journals with more than 100 cited references for the period 1990–2015

management. It is apparent from these results is that corporate governance research is located in the domain of a number of disciplines.

5.3 Representing information flows

A main argument in this paper is that corporate governance research takes place in multiple disciplines. Interdisciplinary flows are illustrated with a Sankey diagram, a type of flow diagram that shows the links between cited references and their citing article. In a Sankey diagram, the width of the links is proportional to the flow quantity; hence, the wider the link the greater the flow. Information flows in the corporate governance literature by subject are represented visually in the Sankey diagram shown in Figure 4.

Figure 4 shows the subject origin of cited references on the right side and the subject of the journals in which they appeared on the left. The diagram shows dispersion between subject areas and how knowledge moves between disciplines.

The value of the Sankey diagram is the ability to visualize the knowledge dispersion between subject areas. Corporate governance research does not take place in a vacuum; there is cross-fertilization of ideas between subjects that informs the way that the field develops and changes.

While a picture is worth a thousand words, a tabular version of the Sankey data provides additional insight into information flows. Table 3 enumerates both the gross numbers and percentage of citation dispersion.

The patterns of dispersion describe the extent to which research ideas are shared among disciplines that contribute to corporate governance research. Law is an example of an insular discipline – 87% of the cited references in law appear in law journals. At the other end of the spectrum, only 23% of the cited references to economics appear in economics journals. Finance, management, and accounting cite their own articles at rates of 43, 50, and 54% respectively. These results show that corporate governance research takes place in multiple

Rank	Subject	Journal	Cited references	Rank	Subject	Journal	Cited references
1	Finance	J FINANC ECON	27,708	26	Management	J BUSINESS	2,408
2	Finance	J FINANC	20,495	27	Organizational Behavior	J BUSINESS ETHICS	2,304
3	Management	ACAD MANAGE J	10,181	28	Law	HARVARD LAW REV	2,211
4	Management	STRATEGIC MANAGE J	10,145	29	Management	J MANAGE STUD	2,175
5	Accounting	J ACCOUNT ECON	7,766	30	Law	YALE LAW J	2,116
6	Economics	J POLIT ECON	7,678	31	Finance	FINANC MANAGE	2,084
7	Economics	AMER ECON REV	6,694	32	Operations Research	MANAGE SCI	2,066
8	Management	ACAD MANAGE REV	6,574	33	Law	STANFORD LAW REV	2,021
9	Management	ADMIN SCI Q	5,949	34	Organizational Behavior	ORGAN SCI	1,938
10	Accounting	ACCT REV	5,172	35	Economics	ECONOMETRICA	1,913
11	Accounting	J ACCOUNTING RES	4,970	36	Economics	J LAW ECON ORGAN	1,691
12	Economics	Q J ECON	4,603	37	Accounting	CONTEMP ACCT RES	1,660
13	Economics	J LAW ECON	4,221	38	Economics	J ECON PERSPECT	1,633
14	Corp Governance	CORP GOV	3,906	39	Entrepreneurship	J BUSINESS VENTURING	1,579
15	Economics	RAND J ECON	3,364	40	Law	Virginia LAW REV	1,545
16	Finance	REV FINANC STUD	3,336	41	Law	U PENN LAW REV	1,463
17	Law	BUSINESS LAWYER	3,308	42	Entrepreneurship	ENTREP THEORY PRACT	1,434
18	Finance	J CORP FINANC	3,199	43	Law	J CORP LAW	1,424
19	Management	J MANAGE	2,975	44	Organizational Behavior	FAMILY BUSINESS REV	1,339
20	Finance	J BANK FINANC	2,907	45	Sociology	AM J SOC	1,297
21	Finance	J FINANC QUANT ANAL	2,821	46	Accounting	J BUS FINANC ACCOUNT	1,258
22	Management	HARVARD BUS REV	2,817	47	Law	U CHICAGO LAW REV	1,239
23	Intl Business	J INT BUS STUD	2,609	48	Economics	REV ECON STUD	1,226
24	Law	COLUMBIA LAW REV	2,542	49	Sociology	AM SOC REV	1,169
25	Management	J BUS	2,408	50	Economics	J ECON LIT	1,155

Note(s): Percent Totals by Discipline: Finance, 31%; Management 23%; Economics, 17%; Accounting, 10%; Law, 9%; Other Areas: Organizational Behavior, Corporate Governance, Entrepreneurship, International Business, Sociology, and Operations Research, 10%

Information
flows and topic
modeling

Table 2.
Top 50 journals by
number of cited
references

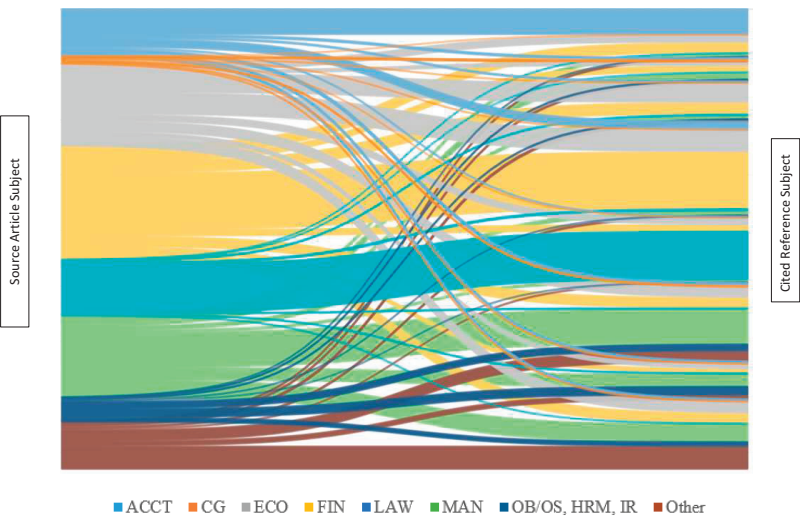


Figure 4.
Source articles and
cited reference
distribution

Journal Subject Code	Cited Reference Subject Code									
		ACCT	CG	ECO	FIN	LAW	MAN	OB	OTHER	Journal Subject Row Total
	ACCT	23,062	409	6,382	9,173	635	2,231	743	1,333	43,968
	Row %	52%	1%	15%	21%	1%	5%	2%	3%	100%
	Column%	54%	6%	9%	9%	1%	3%	3%	3%	
	CG	2,113	2,019	3,034	5,423	585	5,559	1,118	1,453	21,304
	Row %	10%	9%	14%	25%	3%	26%	5%	7%	100%
	Column%	5%	31%	4%	5%	1%	8%	5%	3%	
	ECO	1,493	520	16,950	10,436	1,242	2,622	1,070	2,217	36,550
	Row %	4%	1%	46%	29%	3%	7%	3%	6%	100%
	Column%	4%	8%	23%	10%	2%	4%	5%	5%	
	FIN	7,369	482	19,288	51,320	1,765	3,405	577	1,630	85,836
	Row %	9%	1%	22%	60%	2%	4%	1%	2%	100%
	Column%	17%	7%	26%	50%	4%	5%	3%	4%	
	LAW	1,038	378	5,925	5,619	43,528	1,767	510	918	59,683
	Row %	2%	1%	10%	9%	73%	3%	1%	2%	100%
	Column%	2%	6%	8%	5%	87%	2%	2%	2%	
	MAN	3,074	1,310	8,871	9,190	1,030	31,494	6,171	9,965	71,105
	Row %	4%	2%	12%	13%	1%	44%	9%	14%	100%
	Column%	7%	20%	12%	9%	2%	43%	28%	23%	
	OB	2,216	741	3,778	3,432	893	11,015	7,805	4,359	34,239
	Row %	6%	2%	11%	10%	3%	32%	23%	13%	100%
	Column%	5%	11%	5%	3%	2%	15%	36%	10%	
	OTHER	2,096	740	9,599	8,994	600	15,917	3,965	20,708	62,619
	Row %	3%	1%	15%	14%	1%	25%	6%	33%	100%
	Column%	5%	11%	13%	9%	1%	22%	18%	49%	
	Cited Reference Column Total	42,461	6,599	73,827	103,587	50,278	74,010	21,959	42,583	415,304
	Column %	100%	100%	100%	100%	100%	100%	100%	100%	

Table 3.
Subject Dyad table

disciplines and that researchers frequently cite articles outside of their “home” discipline in their research.

5.4 Latent Dirichlet allocation results

The LDA results show yet another way of describing research areas in corporate governance. The difference between citation analysis and LDA is the ability, using machine learning, to uncover aspects of corporate governance research that are not readily apparent using citation analysis. The advantage of LDA is the ability to extract from the corpus of articles information about topics that are not readily apparent by just reviewing the articles themselves.

LDA requires researchers to select the number of topics desired in the results before analysis begins. We selected a twenty topic model for our LDA analysis, and the results provided twenty numbered topics without names. Labels for the topics were assigned based on the top twenty words generated by the LDAvis software. The topics in the table are arranged in descending order of articles allocated to that topic. Results for each topic also include a list of the top five subject areas where articles on that topic are found, a concentration ratio for the top five subject areas, a list of the top 10 journals associated with the topic, and the twenty terms generated by the LDA software are shown in [Table 4](#).

The LDA results are interesting for a number of reasons. They suggest reasons why finance articles are heavily represented in the retrieved articles. There are eight topics (1, 4, 9, 10, 12, 14, 15 and 20) where finance journals range between 28 and 65% of the journals in that category. The topics include the “Effect of corporate governance in firm performance,” “Executive compensation,” “Boards of directors,” and “Audit committees” and show the sweep of topics covered by finance journals. Among the disciplines represented in the LDA topic journal listings, finance appears in 17 topics with a median rank of 1; economics appears in 16 topics with a median rank of 2, and management appears in 15 topics with a median rank of 3.5.

LDA analysis is useful for uncovering areas of emphasis in the corpus of articles that might not be discovered with regular bibliographic analysis. The subject distribution of the articles described using bibliometric analysis is straightforward. LDA goes beyond citation analysis to uncover topics within the corpus that may be overlooked with traditional methods of analysis.

There are similarities between the LDA results and the keywords used to search Web of Science. For example, agency theory and executive compensation are both topics in the LDA results. Corporate governance is reflected in two of the LDA topics – “Effect of corporate governance in firm performance [Topic 1]” and “Corporate governance theory [Topic 5],” reflecting a bifurcation between theoretical and practical approaches to corporate governance in the literature. LDA topics also include two topics related to family firms (Organizational theory in family firms [Topic 18] and Control of family firms [Topic 9]), two related to boards of directors (Boards of directors [Topic 10] and Fiduciary duty of corporate directors [Topic 19]), and two related to executive compensation (Executive compensation [Topic 15] and Executive compensation/CEO incentive pay [Topic 3]). Beyond the topics already mentioned, there are seven topics related to finance or legal issues [Topics 2, 4, 6, 7, 12, 14, 17]. The preponderance of topics related to financial and legal issues, and the heavy reliance on finance, economics and legal literature, demonstrates that corporate governance research has interdisciplinary breadth.

6. Discussion

6.1 Implications for corporate governance researchers

This research provides an empirically based illustration of the interdisciplinary nature of corporate governance research. To our knowledge, this is the first time citation analysis,

Table 4.
Latent Dirichlet
allocation (LDA) topic
results

Topic number	1	5	18	9	15					
Label	Effect of corporate governance on firm performance	Corporate governance theory	Organizational theory in family firms	Control of family firms	Executive compensation					
Total articles	1720	1,434	1,179	1,043	794					
	Topic	N	Percent	Topic	N	Percent	Topic	N	Percent	
Top 5 subjects by article count	FIN	475	27.6	MAN	416	35.3	FIN	263	33.1	
	ECO	299	17.4	OB/OS	161	13.7	ECO	162	15.5	
	CG	210	12.2	ENTREP	142	12.0	MAN	101	9.7	
	MAN	186	10.8	IB	113	9.6	CG	76	6.4	
	ACCT	156	8.1	MAR	80	6.8	ACCT	60	5.8	
Top 5, % tot articles			77.1			77.3			82.2	87.9
Top-10 journals	CORP GOV	J BUS ETHICS	STRATEGIC MANAGE J	J CORP FINANC	J FINANC ECON					
	J CORP FINANC	CORP GOV	ACAD MANAGE J	J BANK FINANC	J CORP FINANC					
	J BANK FINANC	J MANAGE STUD	J MANAGE	J FINANC ECON	J ACCOUNT ECON					
	CORP GOV-OXFORD	BUS ETHICS Q	J MANAGE STUD	CORP GOV	ACCOUNT REV					
	J BUS ETHICS	ACAD MANAGE REV	ENTREP THEORY PRACT	J FINANC	J FINANC					
	J FINANC ECON	BUS SOC	INT J HUM RESOUR MAN	J FINANC QUANT ANAL	ACAD MANAGE J					
	J BUS RES	ACCOUNT ORG SOC	FAM BUS REV	STRATEGIC MANAGE J	STRATEGIC MANAGE J					
	EMERG MARK FINANC TR	CORP GOV-OXFORD	J BUS RES	J BUS RES	REV FINANC STUD					
	AFR J BUS MANAGE	ORGAN SCI	J BUS VENTURING	PAC-BASIN FINANC J	MANAGE SCI					
	(continued)									

Topic number	1	5	18	9	15
Label	Effect of corporate governance on firm performance	Corporate governance theory	Organizational theory in family firms	Control of family firms	Executive compensation
Total articles	1,720	1,434	1,179	1,043	794
	Topic	N	Percent	Topic	N
	ACCOUNT FINANC	MANAGE DECIS	J INT BUS STUD	EMERG MARK FINANC TR	CONTEMP ACCOUNT RES
Total journals	267	293	216	174	181
Words	governance corporate firms China Chinese disclosure study elsevier listed reserved paper mechanisms performance companies practices political results findings market government	social csr theory stakeholder stakeholders ethical ethics environmental responsibility organizations management organization corporate research moral issues governance organizational society practices	family research strategic innovation organizational performance theory study resources business resource knowledge firms firm strategy relationships relationship orientation influence management	ownership firms shareholders family rights firm controlling shareholder cash structure flow owners concentration shares rd large value voting minority	compensation pay ceo ceos performance incentives executive options incentive risk stock managerial measures executives rms managers bonus salary measure option

(continued)

Information
flows and topic
modeling

Table 4.

Table 4.

Topic number	10	20	7	4	11
Label	Boards of directors	Market returns for IPOs	Audit committees	Credit risks of firm financing	International business
Total articles	662	619	584	449	415
	Topic N Percent	Topic N Percent	Topic N Percent	Topic N Percent	Topic N Percent
<i>Top 5 subjects by article count</i>	FIN 187 MAN 161 CG 81 ACCT 63 OB/OS 59	FIN 404 ACCT 66 ECO 50 MAN 38 IB 17	ACCT 367 FIN 115 MAN 27 OB/OS 22 CG 19	FIN 255 ECO 89 ACCT 30 LAW 15 MAN 11	ECO 136 BHist 60 MAN 43 IB 39 LAW 37
Top 5, % tot articles	28.2 24.3 12.2 9.5 8.9	65.3 10.7 8.1 6.1 2.7	62.8 19.7 4.6 3.8 3.3	56.8 19.8 6.7 3.3 2.4	32.8 14.5 10.4 9.4 8.9
	83.2	92.9	94.2	89.1	76.0
<i>Top-10 journals</i>	CORP GOV J CORP FINANC J FINANC ECON J BUS ETHICS ACAD MANAGE J STRATEGIC MANAGE J J FINANC FINANC MANAGE	J FINANC ECON J CORP FINANC J BANK FINANC J FINANC REV FINANC STUD FINANC MANAGE J FINANC QUANT ANAL J ACCOUNT ECON	AUDITING-J PRACT TH ACCOUNT REV CONTEMP ACCOUNT RES J ACCOUNT ECON J BUS FINANC ACCOUNT J ACCOUNT PUBLIC POL ACCOUNT HORIZ J BUS ETHICS	J BANK FINANC J FINANC ECON J CORP FINANC J FINANC J FINANC QUANT ANAL REV FINANC STUD J FINANC INTERMED J BUS FINANC ACCOUNT	BUS HIST CORP GOV BUS HIST REV ENTERP SOC ADMIN SCI QUART Asia PAC J MANAG ECON GEOGR ECON SOC

(continued)

Topic number	10	20	7	4	11
Label	Boards of directors	Market returns for IPOs	Audit committees	Credit risks of firm	International business
Total articles	Topic N Percent	Topic N Percent	Topic N Percent	Topic N Percent	Topic N Percent
	662	619	584	449	415
	J MANAGE STUD	ACCOUNT REV	ACCOUNT FINANC	J FINANC SERV RES	REV INT POLIT ECON
	J BANK FINANC	J BUS FINAN	REV ACCOUNT STUD	FINANC MANAGE	SOCIO-ECON REV
Total journals					
Words	149	104	82	275	151
	board	firms	audit	debt	institutional
	directors	year	earnings	bank	network
	boards	sample	reporting	banks	japanese
	ceo	returns	accounting	credit	institutions
	director	table	auditor	loan	networks
	outside	firm	auditors	financing	japan
	independent	variable	financial	leverage	german
	independence	variables	accruals	loans	countries
	composition	results	committee	risk	companies
	members	panel	quality	banking	national
	committee	median	sox	equity	pension
	size	coefficient	management	financial	germany
	nonexecutive	value	fraud	crisis	banks
	ceos	period	internal	capital	corporate
	performance	significant	forecast	firms	capitalism
	companies	market	fund	assets	political
	committees	return	analysts	default	world
	executive	ipo	analyst	rating	international
	duality	regression	fees	maturity	business
	monitoring	total	auditing	lending	industrial

(continued)

Information
flows and topic
modeling

Table 4.

Topic number	6			14			2			19			8		
Label	Shareholder legal issues			Protection of investors in foreign firms			Venture capital investments			Fiduciary duty of corporate directors			Agency theory		
Total articles	394	Percent	Topic	341	Percent	Topic	338	Percent	Topic	283	Percent	Topic	263	Percent	Topic
Top 5 subjects by article count	LAW	214	54.3	FIN	169	49.6	ECO	96	28.4	LAW	238	84.1	ECO	134	51
	ECO	57	14.5	ACCT	55	16.1	MAN	44	13	ECO	12	4.2	FIN	53	20.2
	FIN	49	12.4	ECO	53	15.5	ORMS	32	9.5	OB/OS	11	3.9	ORMS	28	10.6
	CG	20	5.1	IB	31	9.1	FIN	30	8.9	FIN	6	2.1	ACCT	17	6.5
	ACCT	17	4.3	MAN	9	2.6	LAW	29	8.6	MAN	4	1.4	MAR	10	3.8
Top 5, % tot articles	90.6			93.0			68.3			95.8			92.0		
Top-10 journals	EUR BUS ORGAN LAW REV			J FINANC ECON			BUS HIST			BUS LAWYER			MANAGE SCI		
	BUS LAWYER			J BANK FINANC			LONG RANGE PLANN			U ILLINOIS LAW REV			J FINANC		
	CORP GOV			J FINANC			INT J HUM RESOUR MAN			SECUR REGUL LAW J			REV FINANC STUD		
	J FINANC ECON J BUS ETHICS			REV FINANC STUD J CORP FINANC			J FINANC ECON SERV IND J			WASH LAW REV VANDERBILT LAW REV			ACCOUNT REV ECON THEOR		
	J CORP FINANC SECUR REGUL LAW J			J INT BUS STUD J BUS FINANC ACCOUNT			CORP GOV HARVARD BUS REV			J BUS ETHICS			ECON LETT EUR J OPER RES		
	STANFORD LAW REV			EUR FINANC MANAG			IND CORP CHANGE			COLUMBIA LAW REV			J ECON BEHAV ORGAN		

(continued)

Topic number	6	14	2	19	8
Label	Shareholder legal issues	Protection of investors	Venture capital	Fiduciary duty of	Agency theory
Total articles	394	341	338	283	263
	Topic N Percent	Topic N Percent	Topic N Percent	Topic N Percent	Topic N Percent
Total journals	U CHICAGO LAW REV	CORP GOV-OXFORD	RES TECHNOL	HASTINGS LAW J	EUR ECON REV
	U PENN LAW REV	FINANC MANAGE	J BUS VENTURING	MICH LAW REV	J ECON
Words	135	104	186	88	117
	law shareholders securities rules regulation voting shareholder corporate legal vote states public company regulatory market united investor takeover disclosure state	countries country foreign protection investor ifrs firms legal vc listing crosslisting domestic investors crosslisted index variables institutional institutions origin distance	venture capital new product market products markets technology exit competition industry investments production food competitive economy development capitalists software customers	corporation directors duty supra law court note fiduciary courts id care corporate judgment delaware litigation liability rule faith rev del	optimal contract agent equilibrium effort manager agents cid cost proposition model contracts utility costs principal problem assume information constraint expected

(continued)

Information
flows and topic
modeling

Table 4.

Topic number	16	17	12	13	3
Label	State-owned firms	franchising	insurance risk	Gender issues in the supply chain	Executive compensation/ CEO Incentive Pay
Total articles	233	140	74	52	40
	Topic N Percent	Topic N Percent	Topic N Percent	Topic N Percent	Topic N Percent
Total journals	CORP GOV	PUBLIC RELAT REV	J BUS FINAN ACCOUNT	PROJ MANAG J	EUR ACCOUNT REV
	113	67	34	36	27
Words	privatization	franchising	insurance	women	tax
	efficiency	franchise	insurers	project	stock
	government	bankruptcy	premiums	female	income
	restructuring	dip	leadership	gender	plan
	enterprises	franchisors	mutual	mm	option
	labour	franchisees	insurer	outsourcing	executive
	soes	outlets	star	hr	options
	productivity	advertising	life	supplier	plans
	privatized	franchised	mutuals	male	exercise
	employee	creditors	policyholders	men	employer
	private	reorganization	transformational	projects	deferred
	state	trustee	guaranty	suppliers	taxable
	public	debtor	underwriting	sponsorship	reg
	enterprise	chapter	property/liability	clan	executives
	transition	franchisor	leaders	goodwill	irc
	employment	franchisee	followers	thrifts	treas
	plants	film	risk	sm	compensation
	sector	debtors	reinsurance	vendor	who
	union	unsecured	nonlife	services	section
	stateowned	br	policyholder	womens	cid

Information
flows and topic
modeling

Table 4.

LDA, and Sankey diagrams have been used jointly in exploring this interdisciplinary subject. This set of tools allowed us to reveal the intellectual genealogy of the field, as well as identify the current topics and specific directions the overarching field is taking. These tools are beneficial because of their power and scope, and because they make no assumptions about the underlying structure of literature being examined. Not only did we identify top journals and articles in a traditional citation networking sense, but also we have been able to trace content in terms of topics and disciplines back to their sources.

This research goes beyond citation analysis by revealing the full intellectual structure of the governance field. Citation analysis is the traditional method for exploring academic networks. LDA enhances this exploration because it frees the analysis from the assumptions of particular sub-fields. By capturing the complete picture of the topical components of a field, LDA provides a means for investigating how different journals deal with specific topics in an interdisciplinary area. It uncovers the topics emblematic of various meta-disciplines. In our research, we have used it to connect journals with related topics. Being able to look at source journals for each topic gives researchers a sense for the topics that come primarily from specific disciplines. The Sankey diagram is another powerful tool for visually displaying the nature of interdisciplinary research areas.

Our results clearly show that the majority of corporate governance research is being published in the fields of accounting, economics, and finance. This runs counter to the prevailing wisdom that management researchers dominate the publication of corporate governance scholarship (Chen and Chang, 2010). Indeed, the results found that over 80% of governance research is published outside of the management literature and that governance research predominantly occurred in a number of related academic disciplines. This research may open opportunities for collaboration among corporate governance researchers who previously had little or no knowledge of one another. As the field of governance matures and expands into sub-disciplines, researchers will be inclined to track intellectual histories and indicate long-term sources. This research provides that illustration for future researchers.

This method is transferrable to other disciplines and it follows that this combined method would be effective in understanding other interdisciplinary fields. It provides both a macro and a micro level view of the development of corporate governance research and how it is used by academics and demonstrates that citation analysis and LDA can be used together to explore interdisciplinary subjects.

Our aim with this research was to extend the work of Turnbull (1997), and Durisin and Puzone (2009), by taking their suggestions for expanding the assessment of corporate governance research into interdisciplinary and international arenas. This research shows how the data mining capabilities of LDA allowed the examination of terms in a full corpus of governance research articles, thereby providing granular levels of detail not accomplished in previous literature reviews and citation analyses.

This method succeeded because of collaboration with a data analytics firm in doing the data analysis. This follows the suggestion of Župić and Čater (2015) to partner with information scientists in doing LDA analyses. It illustrates a method for exploring large data sets with LDA and for data-driven research such as that performed at analytics firms. For researchers with aptitude for data gathering and analysis, LDA software is readily available through software repositories such as the University of Massachusetts Amherst (<http://mallet.cs.umass.edu/topics.php>), the Stanford Artificial Intelligence Laboratory (Blei *et al.*, 2003), and Github (<http://www.github.com>).

6.2 Implications for LIS researchers and librarians

Linking the methods of citation analysis and LDA to explore interdisciplinary fields provides exciting possibilities for library and information science researchers. While this paper used corporate governance as an example, the methods described in this paper are generalizable.

There are a number of ways that LIS researchers and librarians can use the methods from this paper to advance LIS scholarship.

Scholarly activity is becoming increasingly specialized. [Alvesson and Sandberg \(2014\)](#) describe this move toward specialization and lament that it leads to parochial thinking in the academy. They recommend strategies for “box-breaking” research that challenges established paradigms in specialized research areas. [Zahra and Newey \(2009\)](#) identify three modes of theory development that move disciplines forward that can succinctly be described as replication, extension, and transformation. They posit that “maximum impact is created when theory building at the intersection uncovers new phenomena that revise the boundaries of existing disciplines and fields while giving birth to new ones” (2009, p. 1059).

Using the methods described in this paper, LIS researchers and librarians have an opportunity to expand basic LIS research and applied research in other interdisciplinary areas. As experts in parsing disciplinary development, LIS researchers can advance LIS scholarship by exploring the diversity of interdisciplinary areas and viewing this exploration as what Alvesson and Sandberg call “box-breaking research.” Exploration of interdisciplinary areas can result in synergies across and between disciplines. Specific areas where this research could be applied in LIS research and practice are detailed below.

6.2.1 Collection development and management. The ability to deconstruct interdisciplinary fields may lead to better collection development decisions. This research can be applied by individual libraries making collection management decisions. It may also be relevant for libraries in consortiums that are making collection management decisions across institutions. With budgets under pressure at many academic institutions, being able to pinpoint subject areas of interest to researchers may help stretch collections funds.

6.2.2 Support for interdisciplinary research. Academic librarians provide support for researchers across the disciplinary spectrum. That support includes collaboration with departmental faculty on grant projects ([Brandenberg et al., 2017](#)), and facilitating cross-disciplinary research ([Williams et al., 2013](#); [Taskin and Aydinoglu, 2015](#)). This paper provides methods librarians can use to contribute to interdisciplinary research efforts.

6.2.3 Information Seeking. There is a strand of literature in LIS research devoted to information seeking behavior. There may be merit in using methods described in this research as another way to explore information seeking. Much of the interdisciplinary information seeking literature is focused on the behavior of researchers at the point of need ([Niu and Hemminger, 2012](#); [Delserone and Dinkelman, 2017](#); [Wellings and Casseldon, 2019](#)). Deconstructing interdisciplinary fields into their component disciplines, then querying researchers about how they assembled those components may provide new insights into the information seeking research process.

6.2.4 Emerging and established research areas. Exploration of emerging research areas can confirm the degree to which they are interdisciplinary. All of the above areas – collections, departmental/faculty support, and information seeking are important research avenues for emerging research areas. For example, [Mryglod et al. \(2016\)](#), explores the evolution of scientific topics related to the Chernobyl nuclear disaster. Additionally, established research areas like gerontology may also benefit from a rigorous examination of their interdisciplinary components.

6.3 Concluding thoughts

Corporate governance is viewed by many management scholars to be a niche area. This research demonstrates that corporate governance contains a richness of interdisciplinary diversity. Knowledge of that diversity may result in greater cross-fertilization among corporate governance researchers in different component disciplines.

The methods in this paper can be replicated in other interdisciplinary fields and they can be applied in a number of LIS research and practice areas.

References

- Acedo, F.J. and Casillas, J.C. (2005), "Current paradigms in the international management field: an author co-citation analysis", *International Business Review*, Vol. 14 No. 5, pp. 619-639, doi: [10.1016/j.ibusrev.2005.05.003](https://doi.org/10.1016/j.ibusrev.2005.05.003).
- Acedo, F.J., Barroso, C. and Galan, J.L. (2006), "The resource-based theory: dissemination and main trends", *Strategic Management Journal*, Vol. 27 No. 7, pp. 621-636, doi: [10.1002/smj.532](https://doi.org/10.1002/smj.532).
- Alvesson, M. and Sandberg, J. (2014), "Habitat and habitus: boxed-in versus box-breaking research", *Organization Studies*, Vol. 35 No. 7, pp. 967-987, doi: [10.1177/0170840614530916](https://doi.org/10.1177/0170840614530916).
- Azar, O.H. and Brock, D.M. (2008), "A citation-based ranking of strategic management journals", *Journal of Economics and Management Strategy*, Vol. 17 No. 3, pp. 781-802, doi: [10.1111/j.1530-9134.2008.00195.x](https://doi.org/10.1111/j.1530-9134.2008.00195.x).
- Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003), "Latent dirichlet allocation", *Journal of Machine Learning Research*, Vol. 3, January, pp. 993-1022.
- Blei, D.M. (2012), "Probabilistic topic models", *Communications of the ACM*, Vol. 55 No. 4, pp. 77-84, doi: [10.1145/2133806.2133826](https://doi.org/10.1145/2133806.2133826).
- Brandenburg, M., Cordell, S., Joque, J., MacEachern, M. and Song, J. (2017), "Interdisciplinary collaboration: librarian involvement in grant projects", *College & Research Libraries*, Vol. 78 No. 3, pp. 272-282, doi: [10.5860/crl.78.3.272](https://doi.org/10.5860/crl.78.3.272).
- Calabretta, G., Durisin, B. and Ogliengo, M. (2011), "Uncovering the intellectual structure of research in business ethics: a journey through the history, the classics, and the pillars of Journal of Business Ethics", *Journal of Business Ethics*, Vol. 104 No. 4, pp. 499-524, doi: [10.1007/s10551-011-0924-8](https://doi.org/10.1007/s10551-011-0924-8).
- Chen, H.P. and Chan, C.H. (2010), "The evolution of corporate governance research", paper presented at the *Finance and Corporate Governance Conference*, Melbourne, available at: <http://research.cgu.edu.tw/ezfiles/14/1014/img/651/99015P.pdf> (accessed 18 December 2019).
- Chen, S. and Wang, Y. (2018), *Latent Dirichlet Allocation*, paper published by the University of California, San Diego, Department of Electrical and Computer Engineering.
- Chiang, C.-H. and Yang, J.-M. (2012), "A bibliometric study of financial risk literature: a historic approach", *Applied Economics*, Vol. 44 No. 22, pp. 2827-2839, doi: [10.1080/00036846.2011.566208](https://doi.org/10.1080/00036846.2011.566208).
- Culnan, M.J. (1986), "The intellectual development of management information systems, 1972-1982: a co-citation analysis", *Management Science*, Vol. 32 No. 2, pp. 156-172, doi: [10.1287/mnsc.32.2.156](https://doi.org/10.1287/mnsc.32.2.156).
- Culnan, M.J. (1987), "Mapping the intellectual structure of MIS, 1980-1985: a co-citation analysis", *MIS Quarterly*, Vol. 11 No. 3, pp. 341-353, doi: [10.2307/248680](https://doi.org/10.2307/248680).
- Delserone, L.M. and Dinkelman, A.L. (2017), "From their mouths to our ears: research practices and needs of agricultural scientists at the University of Nebraska-Lincoln", *Journal of Agricultural & Food Information*, Vol. 18 Nos 3-4, pp. 373-400, doi: [10.1080/10496505.2017.1330691](https://doi.org/10.1080/10496505.2017.1330691).
- Durisin, B. and Puzone, F. (2009), "Maturation of corporate governance research, 1993-2007: an assessment", *Corporate Governance: International Review*, Vol. 17 No. 3, pp. 266-291, doi: [10.1111/j.1467-8683.2009.00739.x](https://doi.org/10.1111/j.1467-8683.2009.00739.x).
- Durisin, B., Calabretta, G. and Parmeggiani, V. (2010), "The intellectual structure of product innovation research: a bibliometric study of the Journal of Product Innovation Management, 1984-2004", *Journal of Product Innovation Management*, Vol. 27 No. 3, pp. 437-451, doi: [10.1111/j.1540-5885.2010.00726.x](https://doi.org/10.1111/j.1540-5885.2010.00726.x).
- Dyer, T., Lang, M. and Stice-Lawrence, L. (2017), "The evolution of 10-K textual disclosure: evidence from latent dirichlet allocation", *Journal of Accounting and Economics*, Vol. 64 Nos 2-3, pp. 221-245, doi: [10.1016/j.jacceco.2017.07.002](https://doi.org/10.1016/j.jacceco.2017.07.002).
- Fernandez-Alles, M. and Ramos-Rodriguez, A. (2009), "Intellectual structure of human resources management research: a bibliometric analysis of the Journal Human Resource Management,

-
- 1985-2005", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 1, pp. 161-175, doi: [10.1002/asi.20947](https://doi.org/10.1002/asi.20947).
- Filatotchev, I. and Boyd, B.K. (2009), "Taking stock of corporate governance research while looking to the future", *Corporate Governance: An International Review*, Vol. 17 No. 3, pp. 257-265, doi: [10.1111/j.1467-8683.2009.00748.x](https://doi.org/10.1111/j.1467-8683.2009.00748.x).
- Georgi, C., Darkow, I.L. and Kotzab, H. (2010), "The intellectual foundation of the Journal of Business Logistics and its evolution between 1978 and 2007", *Journal of Business Logistics*, Vol. 31 No. 2, pp. 63-109, doi: [10.1002/j.2158-1592.2010.tb00143.x](https://doi.org/10.1002/j.2158-1592.2010.tb00143.x).
- Georgi, C., Darkow, I.L. and Kotzab, H. (2013), "Foundations of logistics and supply chain research: a bibliometric analysis of four international journals", *International Journal of Logistics-Research and Applications*, Vol. 16 No. 6, pp. 522-533, doi: [10.1080/13675567.2013.846309](https://doi.org/10.1080/13675567.2013.846309).
- Gundolf, K. and Filser, M. (2013), "Management research and religion: a citation analysis", *Journal of Business Ethics*, Vol. 112 No. 1, pp. 177-185, doi: [10.1007/s10551-012-1240-7](https://doi.org/10.1007/s10551-012-1240-7).
- Harzing, A.W. (2015), *Journal Quality List*, 56th ed., available at: <http://www.harzing.com/jql> (accessed 27 November 2017).
- Hult, G.T.M. (2015), "JAMS 2010-2015: literature themes and intellectual structure", *Journal of the Academy of Marketing Science*, Vol. 43 No. 6, pp. 663-669, doi: [10.1007/s11747-015-0464-3](https://doi.org/10.1007/s11747-015-0464-3).
- Keasey, K., Thompson, S. and Wright, M. (2005), *Corporate Governance: Accountability, Enterprise and International Comparisons*, John Wiley & Sons, New York, NY.
- Koseoglu, M.A. (2016), "Growth and structure of authorship and co-authorship network in the strategic management realm: evidence from the Strategic Management Journal", *Business Research Quarterly*, Vol. 19 No. 3, pp. 153-170, doi: [10.1016/j.brq.2016.02.001](https://doi.org/10.1016/j.brq.2016.02.001).
- Kushkowsky, J.D. (2012), "Charting the growth of entrepreneurship: a citation analysis of FER content, 1981-2008", *Journal of Business & Finance Librarianship*, Vol. 17 No. 3, pp. 201-219, doi: [10.1080/08963568.2012.685035](https://doi.org/10.1080/08963568.2012.685035).
- Kushkowsky, J.D. and Shrader, C.B. (2013), "Developing a core list of journals in an interdisciplinary area: implications for academic faculty and librarians demonstrated in the business literature", *Library Resources & Technical Services*, Vol. 57 No. 1, pp. 51-65, doi: [10.5860/lrts.57n1.51](https://doi.org/10.5860/lrts.57n1.51).
- Moro, S., Cortez, P. and Rita, P. (2015), "Business intelligence in banking: a literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation", *Expert Systems with Applications*, Vol. 42 No. 3, pp. 1314-1324, doi: [10.1016/j.eswa.2014.09.024](https://doi.org/10.1016/j.eswa.2014.09.024).
- Myrglod, O., Holovatch, Yu., Kenna, R. and Berche, V. (2016), "Quantifying the evolution of a scientific topic: reaction of the academic community to the Chernobyl disaster", *Scientometrics*, Vol. 106 No. 3, pp. 1151-1166, doi: [10.1007/s11192-015-1820-2](https://doi.org/10.1007/s11192-015-1820-2).
- Nerur, S.P., Rasheed, A.A. and Natarajan, V. (2008), "The intellectual structure of the strategic management field: an author co-citation analysis", *Strategic Management Journal*, Vol. 29 No. 3, pp. 319-336, doi: [10.1002/smj.659](https://doi.org/10.1002/smj.659).
- Niu, X. and Hemminger, B. (2012), "A study of factors that affect the information-seeking behavior of academic scientists", *Journal of the American Society for Information Science and Technology*, Vol. 63 No. 2, pp. 336-353, doi: [10.1002/asi.21669](https://doi.org/10.1002/asi.21669).
- Podsakoff, P.M., MacKenzie, S.B., Podsakoff, N.P. and Bachrach, D.G. (2008), "Scholarly influence in the field of management: a bibliometric analysis of the determinants of university and author impact in the management literature in the past quarter century", *Journal of Management*, Vol. 34 No. 4, pp. 641-720, doi: [10.1177/0149206308319533](https://doi.org/10.1177/0149206308319533).
- Pugliese, A., Bezemer, P.J., Zattoni, A., Huse, M., Van den Bosch, F.A.J. and Volberda, H.W. (2009), "Boards of directors' contribution to strategy: a literature review and research agenda", *Corporate Governance-an International Review*, Vol. 17 No. 3, pp. 292-306, doi: [10.1111/j.1467-8683.2009.00740.x](https://doi.org/10.1111/j.1467-8683.2009.00740.x).

- Ramos-Rodriguez, A.R. and Ruiz-Navarro, J. (2004), "Changes in the intellectual structure of strategic management research: a bibliometric study of the Strategic Management Journal, 1980-2000", *Strategic Management Journal*, Vol. 25 No. 10, pp. 981-1004, doi: [10.1002/smj.397](https://doi.org/10.1002/smj.397).
- Sagge, S., Sarto, F. and Cuccurullo, C. (2016), "Evolution of the debate on control enhancing mechanisms: a systematic review and bibliometric analysis", *International Journal of Management Reviews*, Vol. 18 No. 4, pp. 417-439, doi: [10.1111/ijmr.12072](https://doi.org/10.1111/ijmr.12072).
- Samiee, S. and Chabowski, B.R. (2012), "Knowledge structure in international marketing: a multi-method bibliometric analysis", *Journal of the Academy of Marketing Science*, Vol. 40 No. 2, pp. 364-386, doi: [10.1007/s11747-011-0296-8](https://doi.org/10.1007/s11747-011-0296-8).
- Schwab, A. and Zhang, Z. (2018), "A new methodological frontier in entrepreneurship research: big data studies", *Entrepreneurship Theory and Practice*, February, pp. 1-12, doi: [10.1177/1042258718760841](https://doi.org/10.1177/1042258718760841).
- Shafique, M. (2013), "Thinking inside the box? Intellectual structure of the knowledge base of innovation research (1988-2008)", *Strategic Management Journal*, Vol. 34 No. 1, pp. 62-93, doi: [10.1002/smj.2002](https://doi.org/10.1002/smj.2002).
- Sharplin, A.D. and Mabry, R.H. (1985), "The relative importance of journals used in management research: an alternative ranking", *Human Relations*, Vol. 38 No. 2, pp. 139-149, doi: [10.1177/001872678503800204](https://doi.org/10.1177/001872678503800204).
- Shiau, W.L. and Dwivedi, Y.K. (2013), "Citation and co-citation analysis to identify core and emerging knowledge in electronic commerce research", *Scientometrics*, Vol. 94 No. 3, pp. 1317-1337, doi: [10.1007/s11192-012-0807-5](https://doi.org/10.1007/s11192-012-0807-5).
- Shleifer, A. and Vishny, R.W. (1997), "A survey of corporate governance", *The Journal of Finance*, Vol. 52 No. 2, pp. 737-783, doi: [10.1111/j.1540-6261.1997.tb04820.x](https://doi.org/10.1111/j.1540-6261.1997.tb04820.x).
- Sievert, C. and Shirley, K.E. (2014), "LDAvis: a method for visualizing and interpreting topics", *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces-Association for Computational Linguistics*, Baltimore, Maryland, June, pp. 63-70, available at: <https://nlp.stanford.edu/events/illvi2014/papers/sievert-illvi2014.pdf>.
- Sugimoto, C.R., Li, D., Russell, T.G., Finlay, S.C. and Ding, Y. (2011), "The shifting sands of disciplinary development: analyzing north American library and information science dissertations using latent dirichlet allocation", *Journal of the American Society for Information Science and Technology*, Vol. 62 No. 1, pp. 185-204, doi: [10.1002/asi.21435](https://doi.org/10.1002/asi.21435).
- Taşkın, Z. and Aydinoglu, A. (2015) "Collaborative interdisciplinary astrobiology research: a bibliometric study of the NASA Astrobiology Institute", *Scientometrics*, Vol. 103 No. 3, pp. 1003-1022, doi: [10.1007/s11192-015-1576-8](https://doi.org/10.1007/s11192-015-1576-8).
- Tihanyi, L., Graffin, S. and George, G. (2014), "Rethinking governance in management research", *Academy of Management Journal*, Vol. 57 No. 6, pp. 535-1543, doi: [10.5465/amj.2014.4006](https://doi.org/10.5465/amj.2014.4006).
- Turnbull, S. (1997), "Corporate governance: its scope, concerns and theories", *Corporate Governance: International Review*, Vol. 5 No. 4, pp. 180-205, doi: [10.1111/1467-8683.00061](https://doi.org/10.1111/1467-8683.00061).
- Van Ees, H., Gabrielsson, J. and Huse, M. (2009), "Toward a behavioral theory of boards and corporate governance", *Corporate Governance: International Review*, Vol. 17 No. 3, pp. 307-319, doi: [10.1111/j.1467-8683.2009.00741.x](https://doi.org/10.1111/j.1467-8683.2009.00741.x).
- Wellings, S. and Casselden, B. (2019) "An exploration into the information-seeking behaviours of engineers and scientists", *Journal of Librarianship and Information Science*, Vol. 51 No. 3, pp. 789-800, doi: [10.1177/0961000617742466](https://doi.org/10.1177/0961000617742466).
- White, H.D. and McCain, K.W. (1998), "Visualizing a discipline: an author co-citation analysis of information science, 1972-1995", *Journal of the American Society for Information Science*, Vol. 49 No. 4, pp. 327-355, doi: [10.1002/\(SICI\)1097-4571\(19980401\)49:4%3C327::AID-ASIA%3E3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-4571(19980401)49:4%3C327::AID-ASIA%3E3.0.CO;2-4).
- Williams, C., J., O'Rourke, M., Eigenbrode, S.D., O'Loughlin, I. and Crowley, S.J. (2013) "Using bibliometrics to support the facilitation of cross-disciplinary communication", *Journal of the*

American Society for Information Science and Technology, Vol. 64 No. 2, pp. 1768-1779, doi: [10.1002/asi.22874](https://doi.org/10.1002/asi.22874).

Yoshikawa, T. and Rasheed, A.A. (2009), "Convergence of corporate governance: critical review and future directions", *Corporate Governance: An International Review*, Vol. 17 No. 3, pp. 388-404, doi: [10.1111/j.1467-8683.2009.00745.x](https://doi.org/10.1111/j.1467-8683.2009.00745.x).

Zahra, S.A. and Newey, L.R. (2009), "Maximizing the impact of organization science: theory-building at the intersection of disciplines and/or fields", *Journal of Management Studies*, Vol. 46 No. 6, pp. 1059-1075, doi: [10.1111/j.1467-6486.2009.00848.x](https://doi.org/10.1111/j.1467-6486.2009.00848.x).

Župič, I. and Čater, T. (2015), "Bibliometric methods in management and organization", *Organizational Research Methods*, Vol. 18 No. 3, pp. 429-472, doi: [10.1177/109442811456262629](https://doi.org/10.1177/109442811456262629).

Supplementary data

Project data files are available at doi: [10.25380/iastate.9941639](https://doi.org/10.25380/iastate.9941639)

Corresponding author

Jeffrey D. Kushkowsky can be contacted at: kushkows@iastate.edu